Tutor

Dr. Jaime Rubio Martínez
Department of Material Science and
Physical Chemistry (UB)

Grau de Química

Treball Final de Grau

Molecular Dynamics with many ligands: Allosteric inhibition of the antiapoptotic Mcl-1 and Bcl-x_L proteins.

Dinámica molecular con múltiples moléculas: Inhibición alostérica de las proteínas antiapoptóticas Mcl-1 y Bcl-x_L.

María Nuria Peralta Moreno

June 2020

Universitat de Barcelona

B:KC Barcelona Knowledge Campus
Campus d'Excel·lència Internacional

Soy de las que piensan que la ciencia tiene una gran belleza. Un científico en su laboratorio no es sólo un técnico: es también un niño colocado ante fenómenos naturales que le impresionan como un cuento de hadas.

Marie Curie

En primer lugar, quisiera agradecer a mi tutor, el Dr. Jaime Rubio Martínez, por sus enseñanzas y consejos, por la simpatía y confianza transmitidas, y sobre todo por dedicarme su tiempo y acompañarme desde el primer momento. Muchas gracias de corazón.

También quisiera dar las gracias a todos los profesores que me han enseñado y acompañado a lo largo de estos años de continuo aprendizaje.

Pero, sobre todo, quisiera dar las gracias a mi familia, especialmente a mis padres y a mi abuelo, que siempre me han apoyado, acompañado y animado a seguir aprendiendo. Os quiero mucho.

REPORT

CONTENTS

1. SUMMARY

The purpose of the present work is the study and selection of allosteric inhibitors of the Mcl-1 and Bcl-x_L, Bcl-2 protein family members, involved in the regulation of the programmed cell death called apoptosis. Given that the over-expression of these antiapoptotic proteins is related to the development of some types of cancer, both proteins are considered potential therapeutic targets for the development of effective cancer treatment. However, the similarity of the Mcl-1 and Bcl-xL active sites leads to undesirable side effects of cancer treatment. Due to the lack of selectivity in compounds that targets the active site of these proteins, the aim of this work is the selection of two drug candidate fragments capable of selectively regulate the apoptotic mechanism of the proteins Mcl-1 and Bcl-x_L by means of their selective inhibition.

For this purpose, six different fragments have been studied in order to predict their binding affinity and selectivity towards both proteins. To accomplish this purpose, four independent Gaussian accelerated Molecular Dynamics and their respective trajectory analysis have been performed for each studied system. Through the evaluation of the obtained results, the identification of the different allosteric binding sites for both proteins has been possible.

As the knowledge of the binding poses is important to determine the correct growth of the fragments, a first proposal of the binding mode has been made. To correctly discern the binding site of each selected fragment, it has been suggested the necessity of extending the study of their binding modes by an increase of the Molecular Dynamics simulation time. Therefore, more precise and reliable results will be obtained to develop larger ligands in further studies.

Keywords: Mcl-1, Bcl-x_L, Allosteric inhibitors, Apoptosis, Cancer, Gaussian accelerated Molecular Dynamics.

2. RESUMEN

El propósito del presente trabajo es el estudio y selección de inhibidores alostéricos de las proteínas Mcl-1 y Bcl-x_L de la familia Bcl-2 involucradas en la regulación de la muerte celular programada, también llamada apoptosis. Puesto que la sobreexpresión de estas proteínas anti-apoptóticas está relacionada con el desarrollo de algunos tipos de cáncer, ambas proteínas se consideran potenciales objetivos terapéuticos (dianas) para el desarrollo de un tratamiento eficaz contra el cáncer. Sin embargo, la similitud entre los centros activos de Mcl-1 y Bcl-x_L puede conducir a efectos secundarios no deseados derivados del tratamiento del cáncer. Debido a la falta de selectividad en compuestos que actúan en el sitio activo de estas proteínas, el objetivo principal de este trabajo consiste en la selección de dos pequeñas moléculas orgánicas capaces de regular selectivamente el mecanismo apoptótico de las proteínas Mcl-1 y Bcl-x_L a través de la inhibición selectiva de las mismas.

Con este propósito, seis fragmentos diferentes han sido estudiados para predecir su selectividad y afinidad de unión respecto a ambas proteínas. Para ello, cuatro dinámicas moleculares aceleradas independientes y sus respectivos análisis de trayectorias han sido realizados para cada sistema estudiado. Mediante la evaluación de los resultados obtenidos, se ha llevado a cabo la identificación de los diferentes sitios de unión alostérica para ambas proteínas.

Los modos de unión de los fragmentos seleccionados como precursores a fármacos han sido estudiados con la intención de establecer el modelo de crecimiento de los mismos. Para poder discernir correctamente el sitio de unión de cada fragmento seleccionado, se ha planteado la necesidad de extender el estudio de su modo de unión mediante un aumento del tiempo de las simulaciones de dinámica molecular llevadas a cabo en este trabajo. De este modo, se obtendrán resultados más precisos y fiables, necesarios para poder obtener ligandos de mayor tamaño en estudios posteriores.

Palabras clave: Mcl-1, Bcl-x_L, Apoptosis, Cáncer, Inhibidores alostéricos, Dinámica Molecular.

3. INTRODUCTION

The integration of computational methods in all the steps of the drug discovery process and the availability of structural information about disease-involved biomolecules have resulted in the computational discovery of successful drugs, even in difficult cases.[1] The computational search of new drug design strategies has emerged from the high incidence of the cancer disease and the elevated number of side effects of the existent treatments.[2] The present work, defined along this line of research, consists in the computational analysis of the interactions between six small molecules and two proteins involved in the deregulation of some types of cancer.

3.1. BIOLOGICAL FUNCTION OF THE MCL-1 AND BCL-X_L PROTEINS

The apoptosis, also known as programmed cell death, is a highly conserved biological process that takes an essential role in the evolution of pluricellular organisms.[3] Apoptosis normally occurs in the development of the immune system, the embryonic development, to maintain cell population in tissues, among others.[4]

Two different pathways can trigger apoptosis, the extrinsic and the intrinsic pathways. The intrinsic or mitochondrial pathway initiates the cell death process when a cellular stress response is detected, whereas the extrinsic or death-receptors pathway is induced from other cells stimuli.[4] Both pathways are caspases activators, that cause an irreversible cell death by the discriminate degradation of the cell's proteins.

The Mcl-1 and Bcl-x_L proteins are antiapoptotic molecules related to the intrinsic pathway of the programmed cell death regulated by the Bcl-2 protein family to which they both belong. These molecules, located in the outer mitochondrial membrane, exert their function by direct binding to proapoptotic proteins to inhibit them. Both molecules have four BH domains homologous to the Bcl-2 protein, the first three domains form a hydrophobic cavity that allows their interaction with BH3-Only proapoptotic proteins while the fourth domain facilitates their insertion in the outer mitochondrial membrane.[5] This hydrophobic cavity, also known as the

active site of the proteins, is highly conserved in the Bcl-2 protein family, and therefore makes their selective inhibition a challenge.

It has been discovered that the overexpression of Mcl-1 or Bcl-x_L leads to the apparition and development of different types of cancer. The Mcl-1 overexpression has been reported in lung, breast, ovarian, prostate and pancreatic cancer,[6] whilst Bcl-x_L overexpression is associated with the acceleration of tumour formation in the brain or spine glial cells, and in breast or colon epithelial cells.[7,8] Consequently, the development of selective drugs against the biological activity of Mcl-1 and Bcl-x_L is critical given the different types of cancer that they induce and the similarity of their active sites.

3.1. DRUG DESIGN

Thanks to the great scientific advances that have taken place in the last century, the study and treatment of diseases has gone from being a trial and error process to be a structured process focused on a molecular point of view. Nowadays, the determination of the molecular pathways involved in a disease and the identification of the target molecule leads to the search and development of novel therapies through a rational drug design.[9]

Essentially, drugs are chemical molecules that interact to a specific biological target with the purpose of modifying its biological activity by inhibition/activation or participation in metabolic processes with the intention of treating a disease or mitigate its effects.[10]

Generally, the market distribution of a new drug needs a long and expensive period of investigation and testing. Typically, the required time from the conception to the approval of a new drug is 10-15 years. During this time, the majority of molecules fail along the way. Is estimated that only one of each 40.000 compounds is finally approved by the Food and Drug Administration (FDA), with an estimated cost of 800 million dollars.[11]

Figure 1. Stages of drug development

For that reason, a huge effort is being made to reduce the amount of time and resources involved. By the identification and characterization of therapeutic targets, through NMR spectroscopy or X-Ray crystallography, and the application of computational tools in the

preliminary stages of the process is possible to identify, select and optimize effective compounds.[12] As a result, only the ones that present the best predicted pharmacological properties will arrive at the clinical testing phase. Thus, the probability of success of the selected molecules in the clinical phase increases, reducing the amount of time and resources needed to develop a new drug.

Up to now, a huge number of target molecules that contribute to the development of diseases had been identified. Consequently, the three-dimensional study of their biomolecular structure has led to the development of what is known as Structure-Based Drug Design (SBDD).[13] Computational tools such as docking are frequently used in SBDD as a first approach to identify compounds that present favourable interactions with the target molecule. SBDD also enables the target binding site identification, and the design of novel ligands to selective inhibit the biomolecule through the desired cavity.

Another method that belongs to the drug design process is the Fragment-Based Drug Discovery (FBDD), which works in the development of drug candidates by the optimization of previously identified fragments that present high binding affinities towards the target biomolecule.

The search and development of more effective cancer treatments is a wide field of application of SBDD and FBDD strategies due to the difficulty of treating this disease and its high incidence among worldwide population.[2]

For that reason, the aim of this project is to determine selective fragments for each target protein (Mcl-1 and Bcl-x_L) through the computational analysis of their protein-ligand interactions.

3.2.1. Fragment-Based Drug Discovery

In the recent past decades, Fragment-Based Drug Discovery (FBDD) has become an extensively used theoretical approach with an important rate of success[14] in the field of drug design. FBDD is a target-based method that enables the identification of starting point molecules and their subsequent optimization to become good drug candidates. Instead of selecting drug candidates from libraries composed of millions of drug-sized compounds, the FBDD starts with reduced collections of smaller compounds named fragments.[15] This method, applied to Molecular Dynamics simulations, also provides a vision of the target molecule binding sites and the binding modes of ligands.

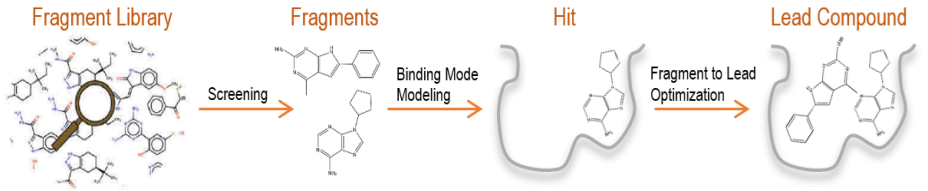

Figure 2. FBDD Process

Fragments are small molecules with pharmacological properties with the ability to develop into larger drug-like molecule denominated as lead compounds.[14] In concordance with Congreve, M et al. the fragments used in FBDD must be selected into conformity with the 'rule of three' (RO3) guideline.[16] This strategy only considers small organic molecules with a molecular weight of less than 300Da, a logP ≤ 3 and a number of hydrogen bond acceptors ≤ 3 and donors ≤ 3 with the object of limit the total number of analysed fragments to those that really have the appropriate characteristics to become a successfully developed drug.

The aim of using small molecules as ligands on the drug design procedure is to facilitate the search of binding sites with its target protein in the absence of ligand-protein complexes structural data and with much more efficiency than larger molecules.[17] Since fragments interact with small protein regions called 'hotspots',[18] is it possible to identify the residues that define the different binding sites of the target molecule.[19] The use of small compounds is crucial in the identification of cryptic binding sites where fragment-specific interactions adapt the conformation of the protein revealing those pockets that are not visible in the isolated form of the protein.

Although the binding interaction between a fragment and a macromolecule may seem weak as a consequence of the entropy energy loss associated with the translation and rotation movements of the fragment in solution, in fact, they form high-quality interactions.[20-22] Estimated as 15-200 kJ/mol and independent to molecular weight,[20,23] this entropic barrier does not vary during the optimization process as a result of the binding mode conservation.[19] Only the fragments that form high quality interactions will be selected to be optimised from fragments to lead compounds with higher affinity and selectivity for the binding site.[24]

Once the protein-ligand binding sites are identified, the selected fragments are typically grown and optimized stepwise to increase their selectivity toward the binding site. This fragment-to-lead process may be undertaken through the merging, linking or growing of the fragment.[19]

The merging method correspond to the combination of the main structural parts and functional groups of both fragments to obtain a larger ligand with more favourable interactions with the binding site than both ligands separately.

Figure 3. Merging method

The linking method consists in the linking of individual fragments, that can bind to different hot spots of the binding site, to obtain a larger ligand that interacts with the whole site, leading to an increase of the binding selectivity of the ligand.

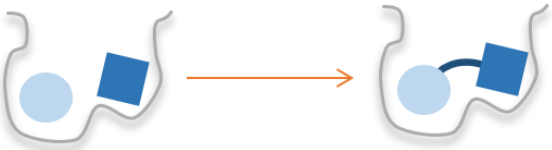

Figure 4. Linking method

The growing method involves the introduction of structural and functional groups to the initial fragment with the purpose of introduce desired properties, interactions and conformation by the selective and gradual "building" of the ligand.

Figure 5. Growing method

The use of FBDD has accelerated the identification of drug candidates. Venetoclax, a Bcl-2 orthosteric inhibitor developed by the Abbott company, and Verburafenib (also known as Zelboraf) an inhibitor of the B-Raf kinase developed by the Plexxikon company, are examples of some approved and commercialized drugs obtained by means of the FBDD method.[14]

4. OBJECTIVES

Given the similarity and high conservation of the Mcl-1 and Bcl-x_L orthosteric binding site, the main objective of this work is the selection of two different compounds that selectively inhibit these proteins through a differentiated allosteric interaction. The fragment-based strategy, selected to accomplish with the purpose of this work, consists on the use of small organic molecules to facilitate the identification of allosteric binding sites in both proteins by the analysis of Molecular Dynamics trajectories.

5. EXPERIMENTAL SECTION

In order to carry out the fragment-based drug discovery process, six small organic molecules were previously selected as fragments that could selectively inhibit the biological targets Mcl-1 and Bcl-x_L. It is important to notice that these fragments were selected considering that they can be experimentally grown by all positions and in concordance with the results of prior studies and the aforementioned 'rule of three'.Through the trajectory analysis of four 400ns Gaussian accelerated Molecular Dynamics (GaMD) for each system and the identification of selective binding sites for both proteins, it is possible to select (for each target molecule) those fragments with higher affinity to an allosteric pocket and make them grow specifically to adapt their conformation to the site to increase their selectivity towards the allosteric site.

For this purpose, the selection of the Mcl-1 and Bcl-x_L selective inhibitors has been carried out by the evaluation of the following descriptors:

- Number of reactive trajectories
- Residence Time (RT)
- K_{DEEP} binding energy (average and best value)
- MMGBSA binding free energy (average and best value)

5.1. MATERIALS AND METHODS

With the purpose of performing the required calculations and evaluation of the obtained results of the present work, the following software has been used:

- AMBER 16 [25]
- Python 3
- MOE 2019.1 *
- The PyMOL Molecular Graphic System
- K_{DEEP} approach (Playmolecule)[26]
- ChemDraw 19.0

(*) All visualizations with MOE had been performed with the agreement/permission of the Chemical Engineering department of the Polytechnic University of Catalonia (UPC).

The scripts used in the present work, provided by the research group of Dr. Jaime Rubio, can be downloaded at https://github.com/DrugDesignUBUJA/fdMD, although a newer version has been used.

The computing unit utilized at the Faculty of Chemistry of the University of Barcelona to fulfil the current work corresponds to a GeForce GTX TITAN GPU with 14 multiprocessors, a core frequency of 0.88 GHz and a global memory size of 6082 MB.

To evaluate the binding affinities of the 6 fragments with the Mcl-1 and Bcl-x_L proteins, the fragment-based Molecular Dynamics (fdMD) method (figure 7) proposed by the Dr. Jaime Rubio's group has been selected.[27] This method offers the advantage of solvating the target protein with already prepared simulation boxes of a solvated ligand since, once prepared, they can be used whenever necessary regardless of the target protein.

With this purpose, six different pre-equilibrated boxes of individual ligands solvated in TIP3P[28] water molecules and the prepared structures of the Mcl-1 and Bcl-x_L proteins, all provided by the research group of Dr. Jaime Rubio, have been used for the construction of the twelve simulation boxes needed to complete the current work. All the system simulation boxes, prepared using the LEaP[29] module of the AMBER software, have been generated by the addition of many solvated ligand boxes considering a box size with a minimum length of any protein atom to the limit of the box of 15 Å and the subsequent removal of any molecule closer than 1.2 Å to the protein in order to avoid water-protein bad contacts (figure 6). Counter ions

needed to neutralize the protein charges are also added. The generalized amber force field 2 (GAFF2)[30] is used to describe the organic molecules and the ff14SB[31] force field to describe the protein. Then, by the use of the ParmEd[32] AMBER module, it is possible to perform the hydrogen mass repartition,[33] to increase the integration time in the molecular dynamics to 0.004ps, and the addition of a Lennard-Jones repulsion terminus to the C99 central atom of each ligand in order to prevent ligand aggregations during the simulations.

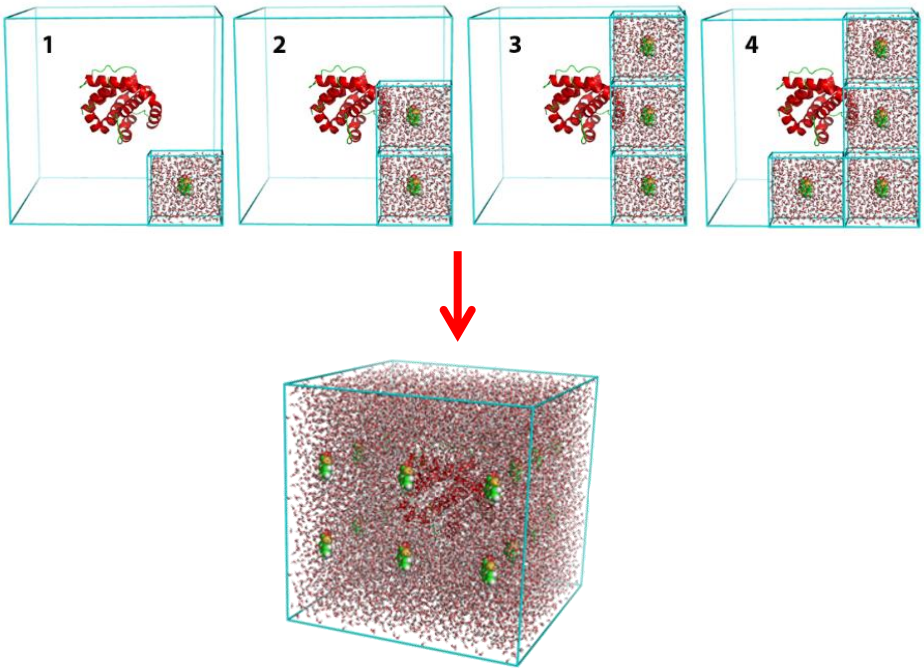

Figure 6. Protein solvation process using many solvated ligand boxes.

Once the system simulation boxes are prepared, they are minimized using 10000 minimization steps of the Steepest Descent method[34] (offered by the Sander module of the AMBER software) to adapt the system to the force field, ensure that the conformation is closely to a minimum of potential energy and to avoid bad orientation of near protein waters. Subsequently, the static simulation boxes are heated to 300 K and equilibrated. Then, four independent Gaussian accelerated Molecular Dynamics simulations (GaMD) of 400 ns are done starting with the same coordinates but using different initial velocities (figure 7, step 4). Finally, the analysis of the trajectories is performed (figure 7, steps 5-6).

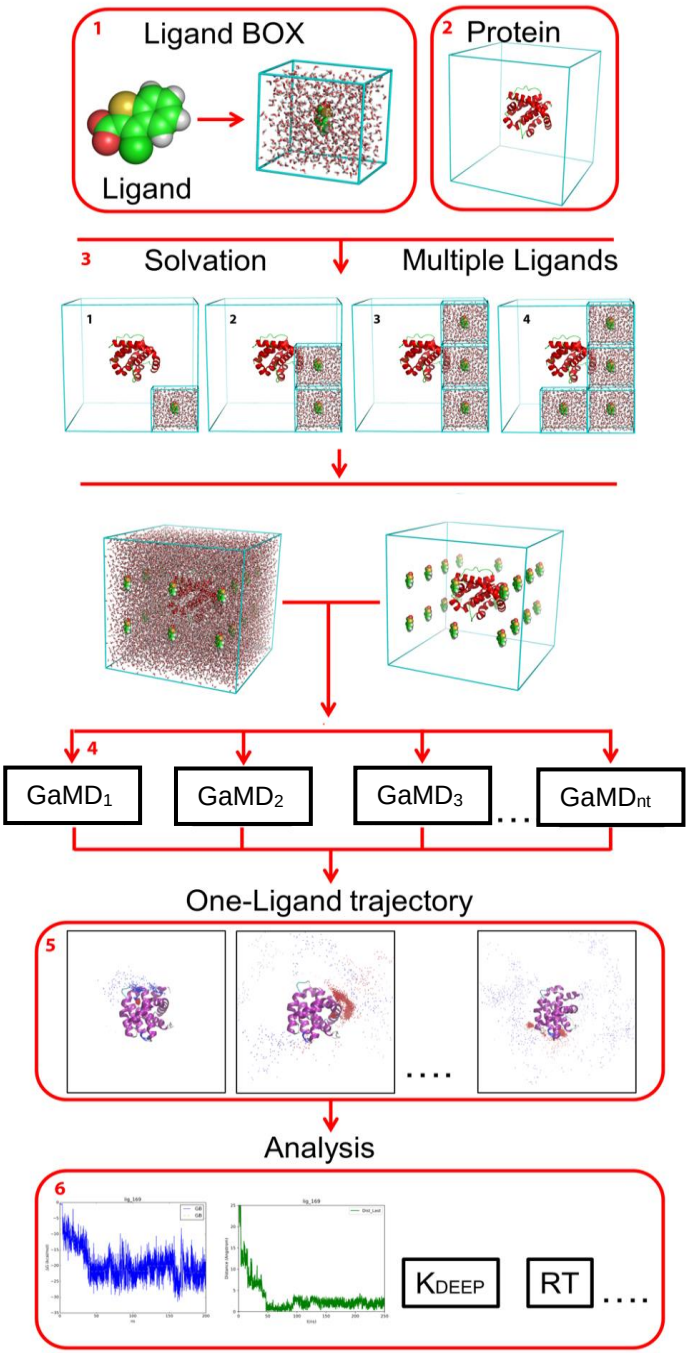

Figure 7. The fdMD method flowchart.

5.2. MOLECULAR DYNAMICS

A Molecular Dynamics (MD) simulation is an extensively used computational technique capable to predict the evolution of a system during an interval of time. By the integration of Newton's equations of motion, this technique can sample the conformational space and the behaviour of a molecular system over time, providing a dynamic model of the studied system[35].

Through the Taylor expansion series for motion (equation 1) the new position of a particle can be calculated considering its initial position, velocity and acceleration.

$$r(t + \Delta t) = r(t) + v(t) \cdot \Delta t + \frac{1}{2} \cdot a \cdot \Delta t^2 \qquad \text{Equation 1}$$

In which the acceleration is calculated from the Second Law of Newton (equation 2) that uses the force determined by the potential energy gradient of the Force Field (equation 3).

$$F_i = m_i \cdot \frac{d^2 x_i}{dt^2} = m_i a_i \qquad F_i = -\frac{dU}{dx_i} \qquad \text{Equations 2 and 3}$$

But, because of the complexity of the potential energy functions, the solution of the Newton's equation can only be achieved through the use of numerical algorithms such as the Verlet algorithm[36] based on third-order Taylor expansions. This method combines the Taylor expansions of the previous and following positions (equation 4 and 5) to easily obtain the new position of a particle in terms of the previous position and the current acceleration (equation 7).

$$r(t + \Delta t) = r(t) + v(t) \cdot \Delta t + \frac{1}{2} \cdot a(t) \cdot \Delta t^2 + \frac{1}{3!} \cdot h(t) \cdot \Delta t^3 \qquad \text{Equation 4}$$

$$+ \quad r(t - \Delta t) = r(t) - v(t) \cdot \Delta t + \frac{1}{2} \cdot a(t) \cdot \Delta t^2 - \frac{1}{3!} \cdot h(t) \cdot \Delta t^3 \qquad \text{Equation 5}$$

$$r(t + \Delta t) + r(t - \Delta t) = 2 \cdot r(t) + a(t) \cdot \Delta t^2 \qquad \text{Equation 6}$$

That can easily be expressed as:

$$r(t + \Delta t) = 2 \cdot r(t) - r(t - \Delta t) + a(t) \cdot \Delta t^2 \qquad \text{Equation 7}$$

This equation is employed during all the simulation procedure to calculate the trajectory of every particle of the system, with the single exception of the initial point, that requires initial velocities assigned by a Maxwell-Boltzmann distribution (equation 8).

$$v_i = \left(\frac{m_i}{2\pi k_B T}\right)^{1/2} \cdot e^{\left(-\frac{m_i \cdot v_i^2}{2 \cdot k_B \cdot T}\right)} \qquad \text{Equation 8}$$

To fulfil any conventional (cMD) or accelerated (aMD) MD simulation, it is advisable to minimize and equilibrate the system to obtain a simulation box as similar as possible to the

physiologic conditions. The importance of minimizing the systems relies on the obtention of the structure of minimum energy adapted to force field and the correction of modelling errors originated from the experimental imprecision of the crystallographic data. The equilibration of the system is an essential step to adapt the system form static to dynamics conditions. Firstly, the system has been submitted to a 200 ps of heating process in which the system has been heated from 0 K to the working temperature of 300 K in 20 ps steps, in all of which the increasement of 30 K and adaptation of the system has taken place. Then, 40 ps of cMD are performed under the NVT Canonical ensemble[37] to adapt the system to the dynamic conditions, maintaining constant the work temperature, the number of molecules and the volume of the simulation box. Secondly, the adaptation of the system density has been done in 60 ps of cMD under the NTP Isothermic-isobaric ensemble,[37] in which the temperature, the number of molecules and the pressure of the simulation box has remained constant and the volume of the simulation box has changed to finally obtain a system density as closer as possible to the water density, avoiding the formation of bubbles. Finally, the last equilibration of the system is performed with 20 ns of cMD under the Canonical ensemble.

All the MD simulations must have been done under the NVT Canonical ensemble and considering Periodic Boundary Conditions (PBC)[38,35] to avoid surface forces by the simulation of an infinite system in which the system box is surrounded by exact system box copies. By doing that, the number of atoms inside the simulation box remains constant because whenever an atom or molecule leaves the simulation box, another exactly the same enters form the opposite side of the simulation box.

Despite the limited simulation length, around the microsecond range, MD simulations are an effective method that enables the study of biological systems of interest. It also permits obtaining information about protein-protein or ligand-protein interactions and the calculous of useful properties such as free binding energies, leading the MD method as an essential step in the drug design field.[39] Sometimes, due to the complexity of some systems and the insufficient simulation time, their study becomes a challenge.

5.2.1. Gaussian accelerated Molecular Dynamics

Whereas conventional Molecular Dynamics (cMD) is unable to represent conformational changes with high potential barriers in such a short period of simulation, Gaussian accelerated

Molecular Dynamics (GaMD) achieves it through the adaptative addition of a boost potential based in harmonic functions and a Gaussian distribution.[40] As a result of applying this boost to

the potential and dihedral energy of the system, the potential barriers decrease obtaining a smoothen energy surface of the system.[41]

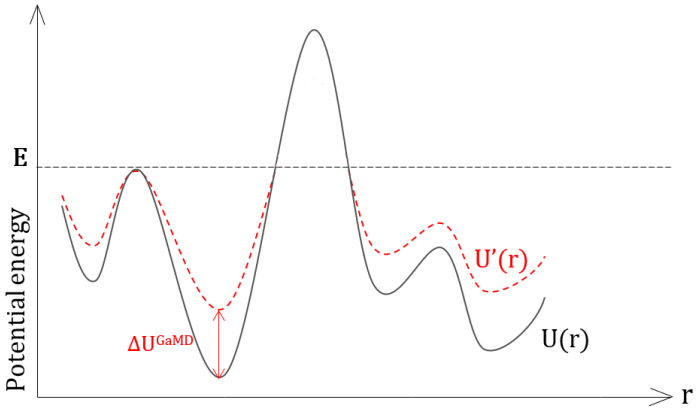

Figure 8. Scheme of the GaMD boosting

As shown in figure 8, the decrease of the potential barriers enables a faster conformational exploration of the studied biomolecule and permits to discard binding sites that present lower binding energies.[42] It has been proved that the results obtained from GaMD simulations are in good agreement with results obtained from highly extended cMD simulations[43] in which some conformational changes cannot be appreciated.

The potential boost provided by the AMBER program, could be defined as the following expression, where k is the harmonic force constant:

$$\Delta U(r) = \frac{1}{2}k(E - U(r))^2 \qquad \text{Equation 9}$$

This expression is applied to the potential energy surface of the system U(r) to obtain the modified system's potential energy, now expressed as:

$$U'(r) = \begin{cases} U(r), & U(r) \geq E \\ U(r) + \Delta U(r), & U(r) < E \end{cases} \qquad \text{Equation 10}$$

Thus, as referred in equations 9 and 10, the addition of the boost potential will only be performed when the potential energy of the system is under the selected threshold energy (E).

The harmonic force constant, utilized in the calculation of the boost potential, changes over time as the system does during the simulation.[41] At first, this parameter is obtained from the last cMD equilibration step of the system simulation box. Then, once the GaMD simulation has started, the recalculation of the harmonic constant is done during all the simulation length.

Using the minimized and equilibrated simulation boxes of the twelve systems as starting points for the simulations, a total of four GaMD independent replicates of 400 ns (100,000,000 steps with an integration step of 0.004 ps in which the equations of motion are solved) have been performed for each studied system.

5.3. TRAJECTORY ANALYSIS

To determine the different reactive binding sites of a targeted protein and rank them according to its stability, it is necessary to perform a complete analysis of the individual ligand trajectories for each simulation.

The residence time, the MMGBSA free binding energy and the K_{DEEP} binding energy values will be considered as a fundamental part of the analysis.

5.3.1. Reactive Trajectories

To discern among reactive and non-reactive trajectories, the Molecular Dynamics of each system, composed by the target molecule and N ligands solvated in water, is stripped into N independent Molecular Dynamics of the protein and a single ligand. Before the division of the global trajectory into independent trajectories (figure 7, step 5), the water molecules and counter ions are removed from the simulation for an easier future analysis.

Once the individual trajectories had been separated, they are all analysed to determine if an interaction between the ligand and the target protein had taken place. This analysis consists of the calculation of the distance between the final pose of the ligand and the rest of the positions taken during the simulation. To do so, the distance calculation is performed with the C99 atom positions, which is designated as the midpoint of the molecule. Only the trajectories that present binding interaction (distance < 5 Å) during at least 90% of the last 20 ns of simulation will be considered as reactive trajectories (figure 10c).

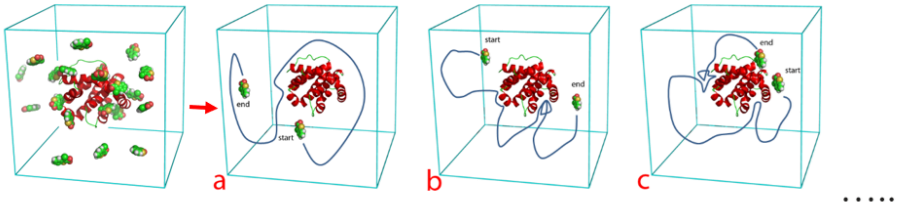

Figure 9. Different types of individual ligand trajectories:
non-reactive trajectories (a and b) and a reactive trajectory (c)

As mentioned above, the determination of the ligand trajectories reactivity relies on the analysis of the final ligand position obtained from the last snapshot of the MD. Only the trajectories whose ligands remain interacting with the protein until the last simulation step of the simulation will be considered as reactive (figure 9c), suggesting the effectiveness of the interaction. Thus, the ones that interact with the protein but do not remain in the pocket will be identified as non-reactive (figure 9b).

5.3.2. Binding Site Identification

Having determined the reactive trajectories, the ligand-protein binding sites were visualized and identified with the MOE software.

With the purpose of defining the reactive binding sites of the target proteins, the amino acids located at less than 4.5 Å from each ligand were identified (see Appendix section). By doing so, 10 and 14 different binding sites were identified for the Mcl-1 and the Bcl-x_L proteins respectively. These binding sites present different types of ligand-protein interactions such as Van der Waals interactions, hydrophobic contacts, hydrogen bonds, aromatic and electrostatic interactions due to the complementarity with some amino acids with the interacting points of all the studied fragments. All the identified binding sites present different binding stabilities and water exposure, for that reason, an evaluation of their binding affinity is required.

5.3.3. Descriptors

At this point, and according with the fdMD method,[26] a consensus of six descriptors are considered to rank the previously identified binding sites. The residence time, the number of reactive trajectories in which a ligand interacts at the analysed binding pocket, the average MMGBSA and K_{DEEP} binding energy and the best MMGBSA and K_{DEEP} binding energy are considered for this purpose.

The residence time (RT) represents the interaction time of a ligand-protein complex. This value, obtained from the previously generated graphs, enables the determination of the time (in ns) that a ligand remains in a possible binding site with a stable interaction. The information given by this descriptor contributes to rank the obtained binding sites depending on their interaction stability. Thus, the estimation of the RT has been calculated for each reactive ligand trajectory and the best value has been considered to rank the different binding sites of each target protein.

With the same purpose, the K_{DEEP} fast machine-learning approach[44] offered by the PlayMolecule website[26] was used to calculate the absolute binding energy of the last snapshot of each trajectory. The importance of considering the K_{DEEP} predictions of the binding energy as a good descriptor relies on the evaluation of 8 different pharmacophoric-like features to determine the protein-ligand interaction affinity.[44] In the current work, only the best K_{DEEP} and the K_{DEEP} average values (equation 11) have been considered to rank the different binding sites.

$$K_{DEEP,\ Average} = \frac{\sum_{i=1}^{nr} K_{DEEP,i}}{4} \qquad \text{Equation 11}$$

The average value consists in the representative K_{DEEP} value assigned for a certain binding site. The equation corresponds to the sum of the K_{DEEP} energy of all the reactive trajectories (nr) divided by the maximum number of reactive trajectories of a pocket, that in the present work corresponds to 4, as 4 is the number of GaMD simulation runs performed.

Another approach utilized to determine binding affinities is the Molecular Mechanics Generalized Born Surface Area (MMGBSA) end-point approximation. This extensively used method[45] enables the free binding energy prediction of biomolecular complexes and, up to now, has been successfully applied to a large number of systems.[46] Due to the problematic of performing the calculations in a solvated medium, in which all the solvent molecules and counter-ions must have been taken into account, the MMGBSA method stablish an implicit solvation model to avoid potential energies fluctuations caused by solvent molecule interactions with the receptor and the ligand during the calculation procedure.[44] Thus, considering the free binding energy as:

$$\Delta G_{binding} = \Delta G^{complex} - \Delta G^{receptor} - \Delta G^{ligand} \qquad \text{Equation 12}$$

A first approach is made:

$$\Delta G_{binding} = \Delta G^{\circ}_{binding} + \Delta G^{\circ\rightarrow sol}_{Complex} - \Delta G^{\circ\rightarrow sol}_{Receptor} - \Delta G^{\circ\rightarrow sol}_{Ligand} \quad \text{Equation 13}$$

Whereby the free binding energy of a component in the vacuum corresponds to the following equation, considering that the internal energy (ΔH°_{i}) is determined by the current force field and the entropy is determined by the vibrational contribution of the Normal Modes of Vibration:

$$\Delta G^{\circ}_{i} = \Delta H^{\circ}_{i} - T \cdot \Delta S^{\circ}_{i} \quad \text{Equation 14}$$

And the solvation energy of a component corresponds to the sum of the electronic and non-polar or hydrophobic contributions:

$$\Delta G^{\circ\rightarrow sol}_{i} = \Delta G^{\circ\rightarrow sol}_{elec} + \Delta G^{\circ\rightarrow sol}_{np} \quad \text{Equation 15}$$

Where the electronic solvation energy is determined by a Generalized Born equation (GB)[46] and the non-polar solvation energy is proportional to the Solvent Accessible Surface Area (SASA):

$$G_{non-polar} = \gamma SASA + \beta \quad \text{Equation 16}$$

The MMGBSA approach is commonly used as an efficient and accurate alternative to the Molecular Mechanics Poisson-Boltzmann Surface Area (MMPBSA) method. The MMGBSA method, utilised in this work, enables a faster prediction of the free binding energy through the use of an easier pair extended equation.[47]

So, in order evaluate the binding stability of the identified binding sites, the free binding energy of the last 20ns of the simulation is considered for each system to calculate their associated descriptors. As in the case of the K_{DEEP} binding energy, the average and the best MMGBSA free binding energy have been considered as descriptors to rank the different binding sites. As stated previously, the maximum number of trajectories corresponds to 4.

$$\Delta G^{MMGBSA}_{average} = \frac{\sum_{i=1}^{nr} \Delta G^{MMGBSA}_{i}}{4} \quad \text{Equation 17}$$

6. BINDING SITES IDENTIFICATION

As previously discussed, the importance of identifying all the possible allosteric binding sites results from the high conservation and similarity of the target proteins active site. Thus, the necessity of finding a fragment that interacts with the target protein in an allosteric site is crucial for the future development of a potential drug candidate.

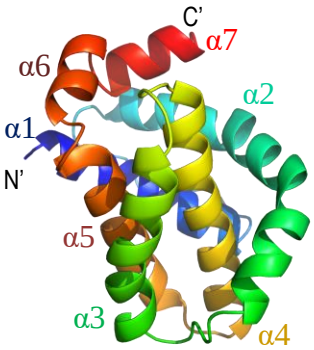

Figure 10. Structural domains of Mcl-1

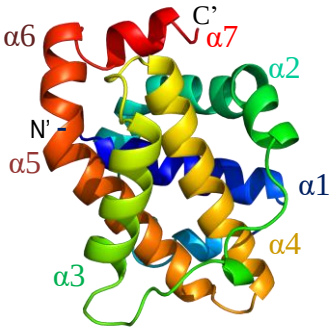

Figure 11. Structural domains of Bcl-x_L

The orthosteric binding site of the Bcl-2 protein family, described by the α2, α3 and α4 domains, is highly conserved. As can be seen in the figure 12, this active site presents four different hotspots named P1, P2, P3 and P4 and an electrostatic interaction in which an arginine is involved.[48,49]

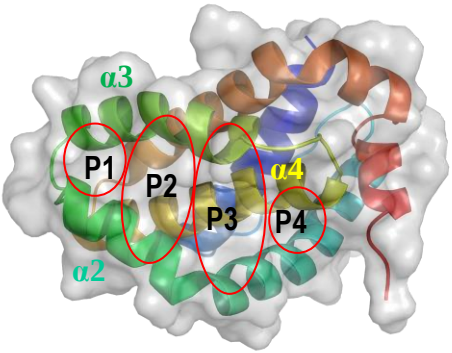

Figure 12. Mcl_1 Orthosteric Site Hotspots

In the present work, most of the ligand interactions with the orthosteric binding site have taken place though the P3 hotspot; although it has been observed simultaneous interactions of two ligands with the P2 and P3 hotspots of the active site (figure 13).

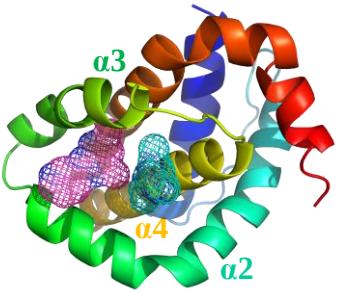

Figure 13. Example of the simultaneous interaction of two p3 fragments in the Mcl-1 orthosteric binding site.

Although the orthosteric binding site has been identified as reactive in both studied proteins, 8 different allosteric binding sites were identified for the Mcl-1 protein in addition to the 12 allosteric binding sites identified for Bcl-x_L.

6.1. MCL-1 ALLOSTERIC BINDING SITES

After carrying out all the experimental process based in the fdMD method, the current study has enabled the determination of different possible allosteric sites that could be structurally analysed in further research as target binding sites of the Mcl-1 protein. By doing that, the development of new drugs could be performed through the SBDD method.

For the Mcl-1 protein, 8 different allosteric binding sites have been found. With the intention of facilitate the identification of the allosteric sites, they had been represented as hotspots over the Mcl-1 secondary structure (Figures 14 and 15).

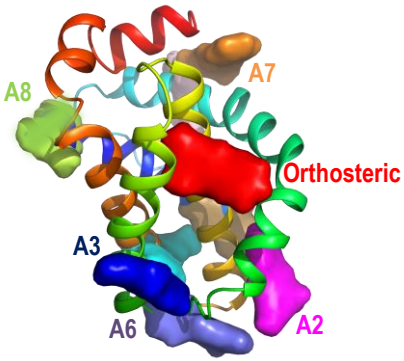

Figure 14. Front view of the Mcl-1 allosteric binding sites.

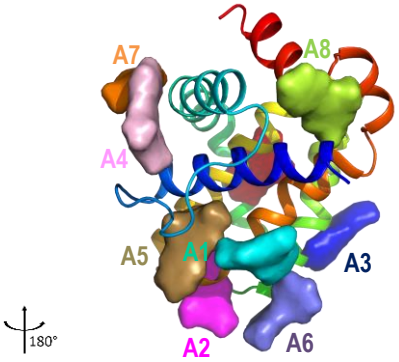

Figure 15. Back view of the Mcl-1 allosteric binding sites.

All the allosteric interactions have been mainly given by hydrogen bonds, Van der Waals, electrostatic and aromatic interactions.

6.2. BCL-x_L ALLOSTERIC BINDING SITES

Regarding to the Bcl-x_L protein, 12 different allosteric binding sites have been identified and represented in the figures 16 and 17. As in the case of Mcl-1, all these allosteric sites could also be studied in further research to develop new allosteric inhibitors for the target protein.

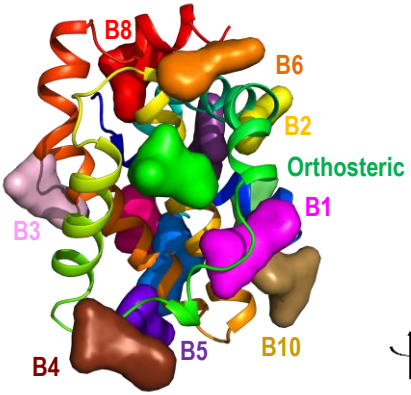

Figure 16. Front view of the Bcl-x_L allosteric binding sites.

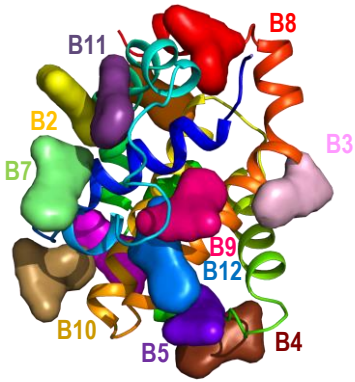

Figure 17. Back view of the Bcl-x_L allosteric binding sites.

As in the Mcl-1 protein, all the allosteric interactions of Bcl-x_L have been mainly given by hydrogen bonds, Van der Waals, electrostatic and aromatic interactions. Considering the similarity between the Bcl-2 protein family members, some of the allosteric binding sites found to the Bcl-x_L may be considered as equivalent to some Mcl-1 allosteric sites. As can be seen in figures 18 and 19, the B5 binding site of the Bcl-x_L protein can be identified as the Mcl-1 equivalent of the A6 binding site.

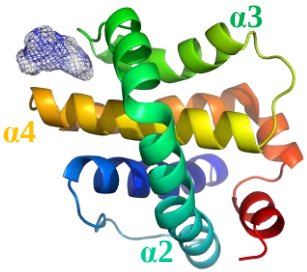

Figure 18. Binding site A6 of the Mcl_1 protein

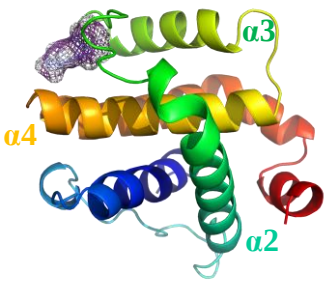

Figure 19. Binding site B5 of the Bcl_x_L protein

7. DRUGGABILITY OF THE ALLOSTERIC POCKETS

Druggable targets are defined as biomolecules whose action can be regulated through their interaction with small molecules. Therefore, the terminus “druggability” designates the ability of certain target molecule binding sites to be modulated by a fragment.[50]

Even though a biomolecule is predicted to be druggable, finding a ligand capable of inhibit it selectively can be a challenge. This may be due to the high conservation of the active site and its high binding competition with other proteins of the same family, as in the case of the Mcl-1 and Bcl-x_L proteins studied in this work. Hence the importance of studying ligands capable of interacting with the protein at different druggable binding sites.

To determine the “druggability” of the previously identified allosteric binding sites of the Mcl-1 and Bcl-x_L proteins, it is important to analyse the results of the descriptors calculations obtained from the trajectory analysis of all the studied systems. Consequently, the obtained results have been tabulated in tables 1 and 2, and their analysis has been done through the evaluation of the descriptors’ best values for each fragment and binding site.

Although the information provided by the descriptors enables the identification of the binding sites in which ligands present better interacting affinities, it is also noteworthy to considerate other factors such as the solvent exposure of the ligand at the binding site and the size of the cavity.

At this point, it is important to clarify that not all the binding sites can be considered as pockets, because some of the allosteric binding sites identified are located over the protein surface and are very exposed to the solvent (figure 20). Thus, only the binding sites that forms a cavity and are not too solvent exposed will be referred as pockets (figure 21). As examples, the B1, B5, B9 and B12 pockets form high quality cavities in which hydrophobic contacts are increased.

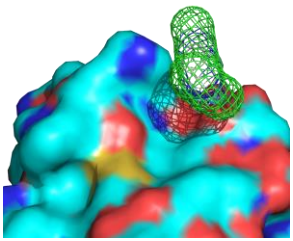

Figure 20. B4 Binding Site of the Bcl-x_L protein

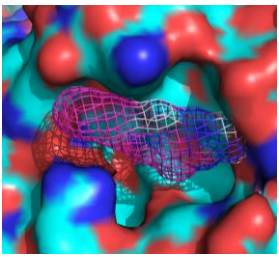

Figure 21. Pocket B1 of the Bcl-x_L protein

	Fragment	Binding site	Ave. MMGBSA[a]	Ave. K_{DEEP}[b]	Best MMGBSA[c]	Best K_{DEEP}[d]	N. react.[e]	Best RT[f]
Mcl-1 protein	p1	Orthosteric Pocket (P3)	**-16.8**	**-5.97**	**-23.5**	**-8.35**	**3**	**390**
		Pocket A2	-7.8	-2.48	-17.9	-5.05	2	150
		Binding Site A3	-6.9	-2.38	-14.1	-5.27	2	160
		Pocket A4	-5.3	-1.34	-21.2	-5.35	1	190
		Pocket A5	-4.7	-1.23	-18.9	-4.93	1	120
		Binding Site A6	-3.9	-0.99	-15.6	-3.95	1	40
		Binding Site A7	-4.1	-1.22	-16.4	-4.88	1	20
	p2	Pocket A1	-4.6	-1.25	-18.4	-4.99	1	50
		Pocket A2	**-10.5**	**-4.08**	-14.3	-5.57	**3**	170
		Pocket A4	-6.2	-1.50	**-24.8**	**-5.98**	1	**250**
		Pocket A8	-5.1	-1.16	-20.4	-4.65	1	60
	p3	Orthosteric Pocket (P3)	**-16.5**	**-5.98**	-20.1	**-7.31**	**4**	220
		Orthosteric Pocket (P2)	-6.4	-1.73	-25.5	-6.90	1	60
		Pocket A2	-3.8	-1.40	-15.0	-5.60	1	20
		Pocket A5	-5.0	-1.28	-20.1	-5.13	1	**320**
		Binding Site A6	-9.4	-2.38	-19.6	-5.70	2	220
		Pocket A8	-6.5	-1.26	**-26.1**	-5.05	1	270
	p4	Orthosteric Pocket (P3)	**-4.9**	**-1.24**	**-19.4**	**-4.97**	**1**	**25**
		Binding Site A7	-3.7	-1.17	-14.9	-4.67	**1**	20
	pd1	Orthosteric Pocket (P3)	**-9.1**	**-3.28**	-18.8	**-7.01**	**2**	**400**
		Pocket A1	-5.4	-1.34	**-21.4**	-5.34	1	60
		Binding Site A3	-6.7	-2.12	-13.9	-4.43	**2**	30
	pd2	Orthosteric Pocket (P3)	-13.3	**-5.34**	-19.9	**-8.25**	**3**	**350**
		Pocket A1	-11.4	-3.46	-19.2	-5.49	**3**	150
		Pocket A4	-7.5	-2.17	-15.9	-4.52	2	190
		Binding Site A6	**-14.0**	-4.30	**-22.6**	-7.56	**3**	170

(a) Average MMGBSA binding energy in Kcal mol^{-1} (see equation 11).
(b) Average K_{DEEP} binding energy in Kcal mol^{-1} (see equation 17).
(c) Best MMGBSA binding energy in Kcal mol^{-1}.
(d) Best K_{DEEP} binding energy in Kcal mol^{-1}.
(e) Number of reactive trajectories of the pocket.
(f) Best Residence Time in ns.

Table 1. Results of the descriptors calculated for all the reactive trajectories of each studied fragment in the Mcl-1 protein.

	Fragment	Binding site	Ave. MMGBSA[a]	Ave. K_{DEEP}[b]	Best MMGBSA[c]	Best K_{DEEP}[d]	N. react.[e]	Best RT[f]
Bcl-x_L Protein	p1	Orthosteric Pocket (P3)	**-17.9**	-5.75	-19.9	-6.37	**4**	270
		Orthosteric Pocket (P2)	-14.5	**-5.88**	-15.0	-6.96	**4**	**380**
		Pocket B1	-17.8	-4.51	**-30.2**	-6.41	3	360
		Pocket B2	-4.9	-1.12	-19.4	-4.46	1	90
		Pocket B3	-3.4	-1.81	-13.7	**-7.24**	1	20
		Binding Site B4	-7.0	-2.71	-14.0	-5.53	2	50
		Pocket B5	-4.6	-1.67	-18.3	-6.67	1	25
		Pocket B6	-4.7	-1.46	-18.8	-5.83	1	60
		Pocket B7	-10.1	-2.23	-22.9	-5.14	2	20
	p2	Orthosteric Pocket (P3)	-14.3	-5.22	-19.7	-6.94	3	80
		Pocket B1	-11.1	-3.17	**-25.2**	**-7.11**	2	100
		Pocket B5	-15.2	-4.34	-24.2	-6.24	3	**240**
		Pocket B6	-9.0	-3.20	-18.6	-7.05	2	30
		Pocket B8	**-18.9**	**-5.68**	**-25.2**	-6.05	**4**	160
		Pocket B9	-16.1	-4.31	-24.1	-6.45	3	150
		Binding Site B10	-3.5	-1.35	-13.8	-5.38	1	25
	p3	Orthosteric Pocket (P3)	**-19.2**	**-6.95**	-23.4	-7.45	**4**	110
		Binding Site B4	-7.5	-3.56	-16.3	**-7.81**	2	65
		Pocket B5	-9.1	-2.92	-18.5	-6.25	2	30
		Pocket B6	-9.2	-3.52	-18.8	-7.08	2	40
		Pocket B9	-2.7	-1.27	-10.9	-5.08	1	20
		Pocket B11	-9.4	-2.41	-19.6	-5.51	2	**160**
		Pocket B12	-6.6	-1.56	**-26.4**	-6.22	1	35
	p4	Orthosteric Pocket (P3)	-3.8	-1.16	-15.0	-4.03	**1**	120
		Orthosteric Pocket (P2)	-3.2	-1.01	-12.9	-4.62	**1**	40
		Pocket B1	**-5.6**	-1.53	**-22.2**	-6.12	**1**	10
		Pocket B6	-3.5	-1.07	-14.1	-4.29	**1**	**130**
		Pocket B8	-5.1	**-1.57**	-20.4	**-6.27**	**1**	30
	pd1	Pocket B2	-3.2	-1.00	-12.8	3.98	1	20
		Binding Site B4	-5.8	**-3.34**	-11.6	**-6.99**	**2**	**110**
		Pocket B6	**-6.4**	-3.18	-14.3	-6.74	**2**	30
		Pocket B8	-4.8	-1.44	**-19.3**	-5.76	1	90
		Binding Site B10	-3.0	-1.29	-11.9	-5.14	1	60
		Pocket B11	-3.4	-1.25	-13.7	-4.99	1	185
	pd2	Pocket B7	**-10.0**	-2.46	**-24.8**	-4.95	**2**	**150**
		Pocket B8	-7.3	**-2.53**	-11.8	-5.54	**2**	100
		Binding Site B10	-3.0	-1.51	-11.8	**-6.05**	1	90

(a) Average MMGBSA binding energy in Kcal mol^{-1} (see equation 11).
(b) Average K_{DEEP} binding energy in Kcal mol^{-1} (see equation 17).
(c) Best MMGBSA binding energy in Kcal mol^{-1}.
(d) Best K_{DEEP} binding energy in Kcal mol^{-1}.
(e) Number of reactive trajectories of the pocket.
(f) Best Residence Time in ns.

Table 2. Results of the descriptors calculated for all the reactive trajectories of each studied fragment in the Bcl-x_L protein.

Considering the results obtained for both proteins, a general predisposition of the ligands to interact at the active site has been observed, to such an extent that simultaneous interactions of two p1 fragments at the same active site of the Bcl-x_L protein have been observed in their four independent GaMD runs. It has also been possible to observe that p2, pd1 and pd2 fragments presents higher affinities to allosteric binding sites rather than for the orthosteric site.

As evidenced in the obtained results, although the p4 fragment has a smaller number of reactive trajectories and simultaneous orthosteric interactions have been observed, it appears to present higher affinities towards the allosteric binding sites of Bcl-x_L protein.

In addition, it has been noted that some allosteric sites appear to be better than others with higher free binding energies. This fact suggests that most of these stable binding sites seem to be less accessible than the apparent best binding site. Therefore, a minor number of ligands can access the allosteric site, so the number of reactive trajectories decreases.

As an example, the p2 fragment interacting in the pocket A4 of the Mcl-1 protein presents better MMGBSA and K_{DEEP} binding energies and residence time than the pocket A2, which presents a major number of reactive trajectories. The higher accessibility of the pocket A2 may be due to its relatively high solvent exposure, in contrast of the pocket A4 which is less accessible and presents a major number of interactions. To clearly determine the preferent binding site of the p2 fragment at the Mcl-1 protein, the GaMD simulations should be extended.

Even though high binding affinities have been found for the B7 and B8 allosteric sites of the Bcl_x_L protein, it is important to note that these binding sites are a guidance and should not be contemplated as druggable binding sites since the 60 residue loop that connects helices α1 and α2 (figure 23), in which both binding sites are located, was found to be very flexible and non-essential for the Bcl_x_L activity[51] so it have not been entirely considered in the Bcl-x_L structure utilized to run the simulations analysed in this work (figure 22).

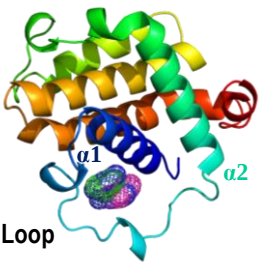

Figure 22. Fragment pd2 at the B7 (green ligand) and B8 (magenta ligand) allosteric binding sites of Bcl-x_L

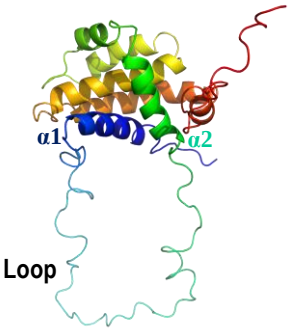

Figure 23. Real Bcl-x_L secondary structure

8. SELECTION OF SELECTIVE FRAGMENTS

The selection of fragments that can selectively regulate the action of the Mcl-1 and Bcl-x_L antiapoptotic proteins through its allosteric inhibition has been undertaken through the evaluation of the obtained results.

The selection process, based on the binding affinity of the fragment with the orthosteric binding site and the information given by the studied descriptors, has enabled the determination of the preferential binding site of each studied fragment.

	Fragment	Most favourable Binding site
Mcl-1 Protein	p1	Orthosteric Pocket (P3)
	p2	**Pocket A2 and A4**
	p3	Orthosteric Pocket (P3)
	p4	Orthosteric Pocket (P3)
	pd1	Orthosteric Pocket (P3)
	pd2	Orthosteric Pocket (P3)
Bcl-x_L Protein	p1	Orthosteric Pocket (P2)
	p2	Pocket B8
	p3	Orthosteric Pocket (P3)
	p4	**Pocket B1** and B8
	pd1	Binding Site B4
	pd2	Pocket B7

Table 3. Most favourable binding sites of the studied fragments

After the analysis, the preferential binding site for each system has been assigned according to the number of best descriptors. The one that presents the maximum number of best descriptors is considered as the most favourable binding site for the system. The other less probable binding sites have not been considered at this stage of the current work.

As stated above, although it has been found that fragments p2, p4 and pd2 present relatively high interaction affinities with the B7 and B8 binding sites of the Bcl-x_L protein, these binding sites should be discarded in the fragment selection process to prevent generating a false drug candidate in further stages of the drug development process, although they could also be studied in future research work to ascertain whether they are good or not.

With regards to the selection of two selective fragments, it is important to consider ligands that present different interactions with both proteins and no affinity towards the orthosteric site of

the target protein to assure its allosteric and selective inhibition. The solvent exposure of the ligand in the pocket must also be considered, since the more water exposed the ligands are, the fewer interactions they can form with the protein.

Another important feature to take into account is the pocket size. The pocket must have the adequate size to permit the increment of the protein-ligand interaction points by growing the ligand. Thus, it is important to locate free protein interaction points nearby the fragment to make it grow selectively to open the cavity and make it more "druggable" as a result of the increased number of interactions.

A reasonable approach to tackle the selective inhibition of Mcl-1 and Bcl-x_L could be the selection of the fragment p2 to selective inhibit the Mcl-1 protein, and the fragment p4 to selective inhibit the Bcl-x_L protein, as the fragment pd2 was preventively discarded to interact with the B7 and B8 binding sites of the Bcl-x_L.

It has been found that the p2 fragment, selected to inhibit the Mcl-1 protein, only exhibit binding affinities for the allosteric binding sites, not presenting any affinity for binding to the orthosteric site of the target protein. In addition, this same fragment present interactions with the orthosteric site of the non-target protein. This fact can be observed in the tables 1 and 2 as well as in figures 24 and 25.

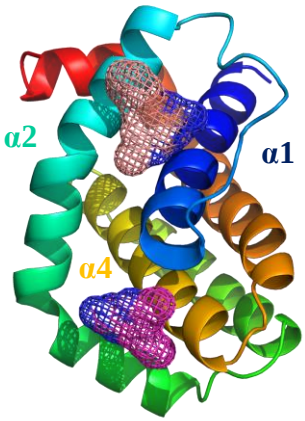

Figure 24. Interaction of the fragment p2 at the A2 (orange) and A4 (pink) allosteric sites of Mcl-1

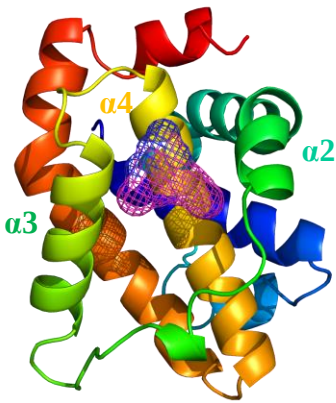

Figure 25. Interaction of the fragment p2 at the Bcl-x_L orthosteric binding site

In the case of the p4 fragment, it presents favourable interactions with the B1 pocket of the Bcl-x_L protein, once discarded the B8 pocket. Although it seems to be less stable due to its small residence time (10 ns), in fact the trajectory of the ligand shows that the fragment p4 remains stable 150 ns in the pre-pocket zone (figure 26) to finally enter inside the pocket cavity (figure 27) and stand stable 10 ns more (figure 28 and 29). For that reason, it might be correct to consider this pocket as a good candidate position to allosteric inhibit the Bcl-x_L protein given that if the dynamics were extended, the most probable is that the ligand would remain stable interacting with the protein inside the pocket.

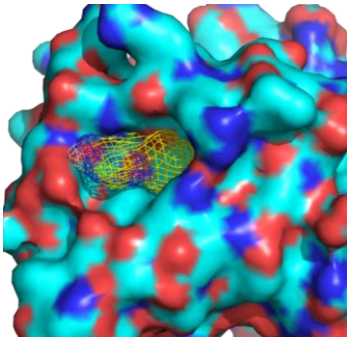

Figure 26. Fragment p4 at the B1 pre-pocket zone of the Bcl-x_L protein

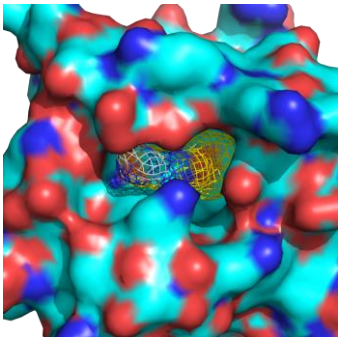

Figure 27. Fragment p4 at the pocket B1 of the Bcl-x_L protein

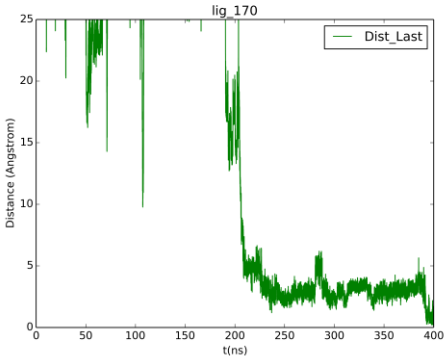

Figure 28. Residence time graph at 25Å of the p4 interaction with the Bcl-x_L protein at the pocket B1

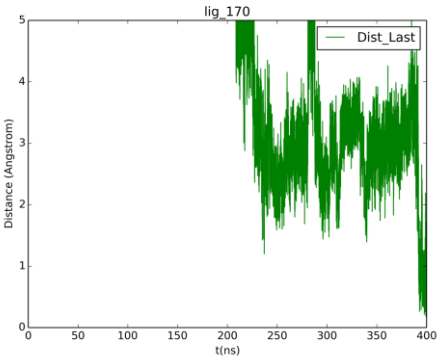

Figure 29. Residence time graph at 5Å of the p4 interaction with the Bcl-x_L protein at the pocket B1

Another option would be the allosteric inhibition of the Bcl-x_L protein by the interaction of the fragment p3 through the pocket B12. This option is not considered in the current work but may be interesting to considerate in further studies.

Once the fragments have been selected, the establishment of their binding mode is essential to determine the interaction-free positions that will be utilized to grow a larger and target-selective ligand. Thus, an exhaustive analysis of their binding mode has been carried out.

For the Mcl-1 protein, it has been noted that 3 of the four GaMD runs ends with the fragment p2 into the A2 allosteric pocket and one of the four dynamic replicates ends in the pocket A4. Superposing the final poses of the reactive ligands over the Mcl-1 protein surface, it has been obtained that p2 fragment's binding mode is preserved at the pocket A2. It has also been observed that the higher MMGBSA and KDEEP energies of the A4 allosteric binding site may be given to the induced fit of the protein to stablish more favourable interactions with the ligand. Another reason, suggested in the previous section, might be its minor solvent exposure comparing to the A2 pocket, which has resulted to be more accessible.

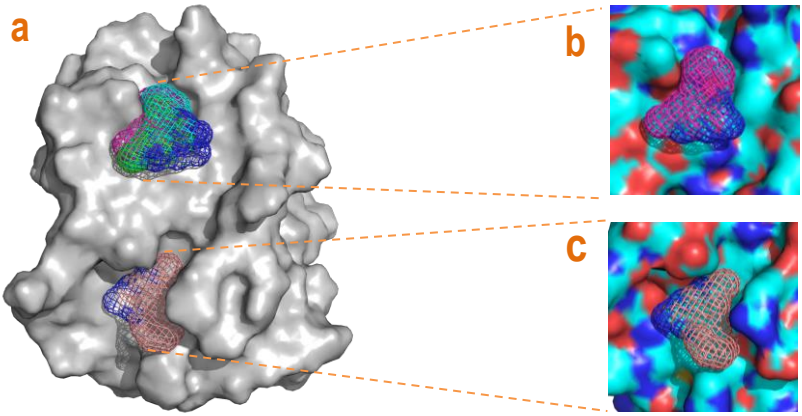

Figure 30. Allosteric inhibition of the Mcl-1 protein (a) and binding mode of the p2 fragment at the A2 (b) and A4 (c) pockets.

Once thoroughly analysed the ligand-protein interactions, it has been possible to determine that the higher binding affinity of the fragment p2 in the pocket A4 is due to electrostatic interaction with the surrounding amino acids and an hydrogen bond established between the ligand (donor) and an aspartic acid of the Mcl-1 structure protein (acceptor), in addition to the Van der Waals interactions that all ligands presents with the protein.

Looking to the figure 31, the binding mode of the fragment in the pocket A4 suggests the molecule growth in the direction of the tyrosine, glutamic and aspartic residues in order to increase the electrostatic interactions with the arginine residue, glutamic and aspartic acids and to stablish another hydrogen bond between the extended NH2 positions of the grown ligand and

the glutamic acid, the tyrosine or the arginine residues. The benzene group of the fragment cannot be used to increase the number of interactions because it is more water exposed than other fragment positions.

Regarding to the A2 pocket (figure 32), the binding mode of the fragment p2 propose the ligand growth in the direction of the benzene that is the less solvent exposed group of the molecule and is located in the direction of a more closed cavity (figure 33b) in which hydrophobic interactions with the leucine residue, electrostatic and hydrogen interactions with histidine and lysine residues can be increased as well as new hydrogen bonds stablished with the lysine, threonine and glutamine residues, and aromatic interactions between fragment and the phenylalanine residue.

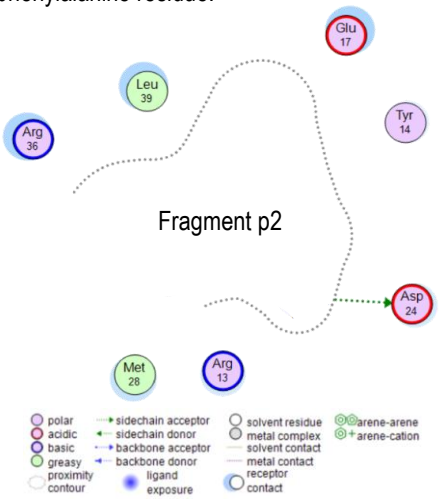

Figure 31. Interactions of the fragment p2 in the pocket A4.

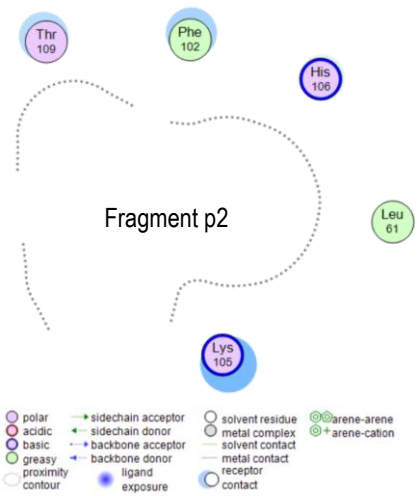

Figure 32. Interactions of the fragment p2 in the pocket A2.

In the case of the Bcl-x_L protein, although it has been observed favourable interactions of the pd2 fragment at the B7 and B8 allosteric binding sites (figure 33), these sites are not representative so the p4 fragment interactions at the B1 allosteric pocket of the Bcl-x_L protein have been thoroughly studied to determine the grown model of the fragment considering its binding mode.

It is important to highlight that, even though the fragment p4 presents less reactive binding sites, it can also be considered as a good candidate because of the high MMGBSA binding energy that the fragment presents interacting at the pocket B1. By the growth of the ligand, its affinity towards the target protein will increase.

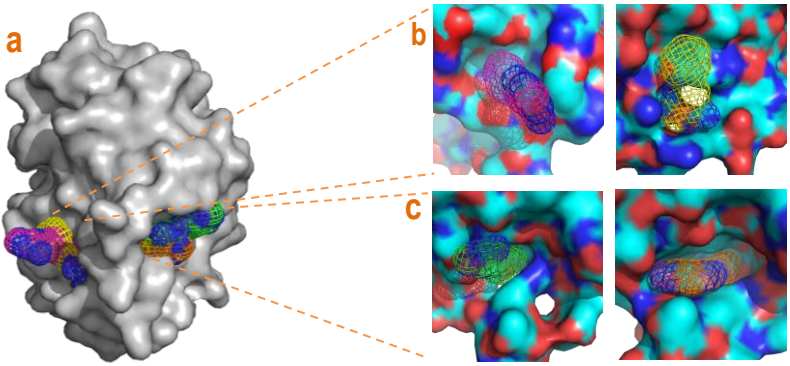

Figure 33. Allosteric inhibition of the Bcl-xL protein (a) and binding modes of the pd2 fragment at the B7 (c) and B8 (b) pockets located at each side of the representative loop.

Although regarding the figure 34 it seems clear that the molecule growth could be done in all directions, it should be preferentially grown in the direction of the lysine, aspartic acid, serine, valine and alanine residues in order to obtain the maximum increase of interactions.

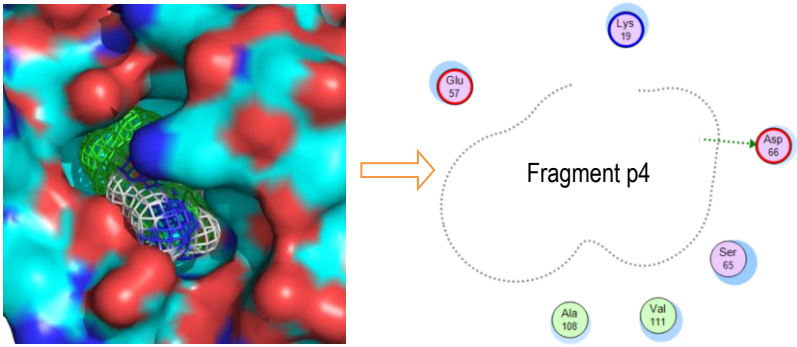

Figure 34. Interactions of the fragment p4 in the pocket B1.

Considering that only exists one representative of the binding mode for the selected fragment, the growth of the molecule cannot be settled until the extension of the GaMD and the subsequent analysis have been undertaken to determine a clear binding mode.

Overall, these results suggest the necessity of extending the simulation length with the aim of obtaining more accurate results that permit to discern the best allosteric binding site of each protein and the clear establishment of the fragments' binding mode. Nevertheless, the selected fragments should be thoroughly studied by means of experimental techniques such as X-Ray spectroscopy or NMR to confirm their binding mode. Then, their growing from hit to lead compounds and optimization could be performed.

9. CONCLUSIONS

Considering that, the main objective of the present work was to select two fragments that could selectively inhibit the Mcl-1 and Bcl-x_L proteins through an allosteric interaction, it has concluded the next points:

- The p2 fragment is a possible candidate to selective inhibit the antiapoptotic function of the Mcl-1 protein through an allosteric interaction.
- The p4 fragment is a possible candidate to selective inhibit the antiapoptotic function of the Bcl-x_L protein through an allosteric interaction.
- To obtain more accurate results, the extension of the GaMD simulation length should be performed. By doing that, the identification of the most favourable binding site of Mcl-1 and the clear determination of the binding modes for each selected fragment will allow the development of better ligands in further studies.

10. REFERENCES AND NOTES

1. Ou-Yang, S. S.; Lu, J. Y.; Kong, X. Q.; et al. Computational drug discovery. *Acta Pharmacologica Sinica*, **2012**, 33, 1131–1140.
2. Chakraborty, S.; et al. The difficulties in cancer treatment. *Ecancermedicalscience*. **2012**, 6, 1-6.
3. Delgado L. *Optimització in silico de compostos antitumorals*. Universitat de Barcelona, **2011**. p 41-46.
4. Elmore, S. Apoptosis: A Review of Programmed Cell Death. *Toxicologic Pathology*, **2007**, 35(4), 495-516.
5. Gurudutta, G. et al. Structural conservation of residues in BH1 and BH2 domains of Bcl-2 family proteins. *FEBS Letters*, **2005**, 579(17), 3503–3507.
6. Xiang, W., Yang, C.-Y., & Bai, L. (2018). OncoTargets and Therapy Dovepress MCL-1 inhibition in cancer treatment. OncoTargets and Therapy, **2018**, 11–7301.
7. Trisciuoglio, D.; Tupone, M. G.; Desideri, M.; et al. BCL-XL overexpression promotes tumor progression-associated properties article. *Cell Death and Disease*, **2017**, 8(12), 3216.
8. Fabregat, I.; Gil, J. Desregulación de la apoptosis en tumores colorrectales y hepáticos. *Gastroenterologia y Hepatologia Continuada*, **2009**, 8(1), 28–31.
9. Delgado L. *Optimització in silico de compostos antitumorals*. Universitat de Barcelona, **2011**, p 12-16.
10. Collin P. *Dictionary of Medical Terms*, 4th ed.; A&C Black: London, **2004**, p 116.
11. Kapetanovic, I. M.; CADD in silico-chemico-bio-aproach. *Chemico-Biological Interactions*, **2009**, 171(2), 165–176.
12. Donald, J. A. *Burger's Medicinal Chemistry & Drug Discovery*, volume 2, 6th ed.; Wiley-Interscience Publication: Virginia, 2003, pp 38-42.
13. Merz, K. M.; Ringe, D.; Reynolds C. H.; *Drug Design Structure and Ligand Based Approaches*, Cambridge University Press: Cambdrige, **2010**, pp 181-196.
14. Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; et al. Twenty years on: The impact of fragments on drug discovery. *Nature Reviews Drug Discovery*, **2016**, 15(9), 605–619.
15. Merz, K. M.; Ringe, D.; Reynolds C. H. *Drug Design Structure and Ligand Based Approaches*, Cambridge University Press: Cambdrige, **2010**, pp 30-40.
16. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A 'Rule of Three' for fragment-based lead discovery? *Drug Discovery Today*, **2003**, 8, 876-877.
17. Jhoti, H.; Williams, G., Rees, D. C.; Murray, C. W. The "rule of three" for fragment-based drug discovery: Where are we now? *Nature Reviews Drug Discovery*, **2013**, 12(8), 644.
18. Ferenczy, G. G.; et al. Thermodynamics of fragment binding. *J Chem Inf Model*. **2012**, 52, 1039–1045.
19. Jacquemard, C.; Kellenberger, E. A bright future for fragment-based drug discovery: what does it hold? *Expert Opinion on Drug Discovery*, **2019**, 14(5), 413–416.
20. Murray, C. W.; Rees, D. C. The rise of fragment-based drug discovery. *Nature Chemistry*, **2009**, 1(3), 187–192.
21. Jencks, W. P. On the attribution and additivity of binding energies. *Proceedings of the National Academy of Sciences of the United States of America.* **1981**, 78, 4046–4050 (1981).
22. Page, M.; Jencks, W. Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. *Proceedings of the National Academy of Sciences of the United States of America.* **1917**, 68, 1678–1683.
23. Finkelstein, A. V.; Janin, J. The price of lost freedom: entropy of bimolecular complex formation. *Protein Eng.* **1989**, 3, 1–3.

24. Kumar, A.; Voet, A.; Zhang, K. Y. J. Fragment Based Drug Design: From Experimental to Computational Approaches. *Current Medicinal Chemistry*, **2012**, 19(30), 5128–5147.
25. Case, D. A.; Betz, R.M.; Cerutti, D. S; et al. AMBER 16, *University of California: San Francisco,* **2016.**
26. PLAYMOLECULE https://playmolecule.org/
27. Cristian Privat; et al. Fragment Dissolved Molecular Dynamics: A systematic and efficient method to locate binding sites. *Journal of Chemical Theory and Computation*, **submitted.**
28. Price, D. J.; Brooks, C. L. A modified TIP3P water potential for simulation with Ewald summation. *J Chem Phys*, **2004**, 121, 10096–10103.
29. Case, D.; et al. *AMBER 2016 Reference Manua*l. University of California: San Francisco, **2016**, p 199.
30. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general Amber force field. *Journal of Computational Chemistry*, **2004**, 25(9), 1157–1174.
31. Maier, J. A.; et al. ff14SB: Improving Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation,* **2015,** 11, 3696-3713.
32. Case, D.; et al. AMBER 2016 Reference Manual. *University of California: San Francisco,* **2016**, p.242.
33. Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *Journal of Chemical Theory and Computation*, **2015**, 11(4), 1864–1874.
34. Frank, J. *Introduction to computational chemistry 2nd ed*. Chichester, Wiley, **2007**, p.383.
35. *Hernández, E. R.* Molecular Dynamics: from basic techniques to applications (A Molecular Dynamics Primer). AIP Conference Proceedings, **2008**, 1077, 95–*123.*
36. Frank, J. *Introduction to computational chemistry, 2nd ed*.; Wiley: Chichester, **2007**, pp 8, 451-453.
37. Alan, H. *Molecular Modelling for Beginners, 2nd ed.;* Wiley: Manchester, **2008,** pp 114-115.
38. Frank, J. *Introduction to computational chemistry, 2nd ed*.; Wiley: Chichester, **2007**, p.464.
39. Karplus, M.; McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nature Structural Biology*, **2002**, 9(9), 646–652.
40. Hospital, A.; Goñi, J. R.; Orozco, M.; Gelpí, J. L. Molecular dynamics simulations: Advances and applications. *Advances and Applications in Bioinformatics and Chemistry*, **2015**, 8(1), 37–47.
41. Miao, Y.; Sinko, W.; Pierce, L.; Bucher, D.; Walker, R. C.; McCammon, J. A. Improved reweighting of accelerated molecular dynamics simulations for free energy calculation. Journal of Chemical Theory and Computation, 201410(7), 2677–2689.
42. Miao, Y.; et al. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. *Journal of Chemical Theory and Computation*, **2015**, 11(8), 3584–3595.
43. Miao, Y.; Feixas, F.; Eun, C.; Mccammon, J. A. Accelerated Molecular Dynamics Simulations of Protein Folding Graphical abstract HHS Public Access. *J Comput Chem.* **2015**, 30(3620), 1536–1549.
44. Jiménez, J.; et Al. KDEEP: Protein-Ligand Absolute Binding Affinity Prediction via 3D-Convolutional Neural Networks. *Journal of Chemical Information and Modeling*, **2018,** 58(2), 287–296.
45. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery*, **2015**, 10(5), 449–461.
46. Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chemical Reviews*, **2019**, 119(16), 9478–9508.
47. Onufriev, A.; Bashford, D.; Case, D. A. Modification of the generalized born model suitable for macromolecules. *Journal of Physical Chemistry B*, **2000**, 104(15), 3712–3720.
48. Denis, C.; Sopková-De Oliveira Santos, J.; Bureau, R.; Voisin-Chiret, A. S. Hot-Spots of Mcl-1 Protein. *Journal of Medicinal Chemistry*, **2020**, 63(3), 928–943.
49. Yang, C. Y.; Wang, S. Analysis of flexibility and hotspots in Bcl-xL and Mcl-1 proteins for the design of selective small-molecule inhibitors. *ACS Medicinal Chemistry Letters*, **2012**, 3(4), 308–312.
50. Owens, J. Determining druggability. *Nature Reviews Drug Discovery*, **2007**, 6(3), 187–187.
51. Muchmore, S. W.; Sattler, M; Liang, H; et al. X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death. *Nature,* **1996**, 381(6580),335-341.

11. ACRONYMS

Bcl-2: B-cell lymphoma - 2

Bcl-xL: B-cell lymphoma - extra large

BH domain: Bcl-2 Homology domain

B-Raf kinase: Serine/Threonine-protein kinase

cMD: Conventional Molecular Dynamics

fdMD: Fragment-dissolved Molecular Dynamics

FBDD: Fragment Based Drug Discovery

FDA: Food and Drug Administration

GAFF2: Generalized Amber Force Field 2

GaMD: Gaussian accelerated Molecular Dynamics

Mcl-1: Myeloid cell leukemia - 1

MD: Molecular Dynamics

MMGBSA: Molecular Mechanics Generalized Born Surface Area

MOE: Molecular Operating Environment (Software)

NMR: Nuclear Magnetic Resonance

NTP: Canonical ensemble

NTV: Isothermal-isobaric ensemble

RO3: Rule of three

RT: Residence time

SBDD: Structure Based Drug Design

TIP3P: Transferable Intermolecular Potential with 3 Points

APPENDICES

APPENDIX 1: DEFINITION OF THE MCL-1 AND BCL-X_L IDENTIFIED POCKETS

Mcl-1

Orthosteric Pocket

	Residue number
Ortho P3	224 227 228 231 249 250 253 254 262 263 264 266 267 270 280
Ortho P2	231 232 234 235 246 247 249 250 267 270 293 294 297 302

Allosteric Pockets

	Residue number
A1	243 244 247 289 290 292 293 296 300
A2	189 190 221 225 229 232 273 274 276 277 279 280
A3	234 235 236 237 242 245 246 248 249
A4	184 185 187 188 193 194 195 197 198 199 207 210 211 214
A5	182 183 184 186 187 190 191 279 284 285 287 288 291
A6	237 240 243 246 277 278 281 283 286 289 290 293
A7	208 209 211 212 215 316 317 319 320 321
A8	172 175 176 179 292

Bcl-X_L

Orthosteric Pocket

Residue number

Ortho P3	97	101	102	104	105	106	108	112	113	126	129	130	132	136	138	139	142	143	145	146	149		
Ortho P2	104	105	108	111	112	122	125	126	127	129	130	142	146										

Allosteric Pockets

Residue number

B1	12	16	19	20	21	94	97	98	101	102	105	106	107	108	109	145	146	148	149	150	152	153	156	157				
B2	11	12	15	34	84	87	88	91	108	109	110	111	112	113	114	115	159											
B3	120	123	124	127	128	131	169	172	173	176	177																	
B4	109	110	111	112	113	114	120	123	124	127	128	129	131	132	133	134	135	168	169	172	173	176	177	178	181	204		
B5	108	109	110	111	112	113	114	115	116	117	118	119	120	122	123	126	150	153	154	159	161	162	165	166	169			
B6	89	90	92	93	94	96	97	100	101	130	137	138	139	141	142	190	191	192	194	195	196							
B7	2	3	4	6	7	10	11	14	18	23	24	25	28	29	30	31	32	33	34	35	36	38						
B8	2	3	4	5	6	7	8	11	32	33	34	39	41	82	83	84	85	86	87	144	168	171	172	174	175	178		
	179	182	183	184	186	187	188	189	192	193																		
B9	2	3	4	6	9	10	13	24	27	28	31	32	33	36	163	164	167	168	171	172	175	176						
B10	17	20	21	22	27	28	29	152	153	155	156	160	161	164	165													
B11	2	3	4	6	7	8	10	11	35	36	39	40	82	83	86	87	188	189										
B12	6	9	10	13	24	27	30	31	32	33	122	123	124	126	168	171												

The highlighted residues are a guidance because the residues from 28 to 41 corresponds to the extremely large loop, that the protein Bcl-xL presents in its real structure connecting alpha-helices 1 and 2. This loop have not been considered in this work due its flexibility and inactivity in the anti-apoptotic protein function.[51]

APPENDIX 2: RESULTS OF MCL-1 AND BCL-XL 400NS FDMD

The results obtained through the trajectory analysis of each studied system are presented in the following tables. The average values have been obtained as in the equations 11 and 17, dividing by four the sum of the values obtained for each GaMD replicate for the same pocket. If a binding site is non-reactive in a dynamic, a zero would be assigned for the descriptors values.

Legend

RT: Residence Time (in ns)

MMGBSA: Molecular Mechanics Generalized Born Surface Area free binding energy (in kcal·mol^{-1})

K_{DEEP}: K_{DEEP} binding energy (in kcal·mol^{-1})

Mcl1 - Fragment p1

	DYN 1			DYN 2			DYN 3			DYN 4		
	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}
Orthosteric P3				390	-23.5	-8.35	290	-21.9	-7.85	160	-21.8	-7.68
Pocket A2				150	-13.2	-4.86				45	-17.9	-5.05
Pocket A3				20	-14.1	-4.24	160	-13.5	-5.27			
Pocket A4				190	-21.2	-5.35						
Pocket A5				120	-18.9	-4.93						
Pocket A6	40	-15.6	-3.95									
Pocket A7										20	-16.4	-4.88

Averages				
Orthosteric P3	210	-16.8	-5.97	3 /4 dyn
Pocket A2	49	-7.8	-2.48	2 /4 dyn
Pocket A3	45	-6.9	-2.38	2 /4 dyn
Pocket A4	48	-5.3	-1.34	1 /4 dyn
Pocket A5	30	-4.7	-1.23	1 /4 dyn
Pocket A6	10	-3.9	-0.99	1 /4 dyn
Pocket A7	5	-4.1	-1.22	1 /4 dyn

Mcl1 - Fragment p2

	DYN 1			DYN 2			DYN 3			DYN 4		
	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}
Pocket A1				50	-18.4	-4.99						
Pocket A2	170	-13.6	-5.44	150	-14.2	-5.31				35	-14.3	-5.57
Pocket A4				250	-24.8	-5.98						
Pocket A8	60	-20.4	-4.63									

Averages				
Pocket A1	13	-4.6	-1.25	1 /4 dyn
Pocket A2	89	-10.5	-4.08	3 /4 dyn
Pocket A4	63	-6.2	-1.50	1 /4 dyn
Pocket A8	15	-5.1	-1.16	1 /4 dyn

Mcl1 - Fragment p3

	DYN 1			DYN 2			DYN 3			DYN 4		
	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}
Orthosteric P3	55	-13.4	-5.80	220	-18.5	-6.41	20	-13.8	-4.25	80	-20.1	-7.31
Orthosteric P2	60	-25.5	-6.90									
Pocket A2	20	-15.0	-5.60									
Pocket A5				320	-20.1	-5.13						
Pocket A6	140	-17.8	-3.81				220	-19.6	-5.70			
Pocket A8	270	-26.1	-5.05									

Averages				
Orthosteric P3	94	-16.5	-5.94	4 /4 dyn
Orthosteric P2	15	-6.4	-1.73	1 /4 dyn
Pocket A2	5	-3.8	-1.40	1 /4 dyn
Pocket A5	80	-5.0	-1.28	1 /4 dyn
Pocket A6	90	-9.4	-2.38	2 /4 dyn
Pocket A8	68	-6.5	-1.26	1 /4 dyn

Mcl1 - Fragment p4

	DYN 1			DYN 2			DYN 3			DYN 4		
	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}
Orthosteric P3										25	-19.4	-4.97
Pocket A7							20	-14.9	-4.69			

Averages				
Orthosteric P3	6	-4.9	-1.24	1 /4 dyn
Pocket A7	5	-3.7	-1.17	1 /4 dyn

Mcl1 - Fragment pd1

	DYN 1			DYN 2			DYN 3			DYN 4		
	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}
Orthosteric P3							400	-18.8	-7.01	400	-17.4	-6.12
Pocket A1				60	-21.4	-5.34						
Pocket A3	30	-13.9	-4.43				20	-12.7	-4.06			

Averages				
Orthosteric P3	200	-9.1	-3.28	2 /4 dyn
Pocket A1	15	-5.4	-1.34	1 /4 dyn
Pocket A3	13	-6.7	-2.12	2 /4 dyn

Mcl1 - Fragment pd2

	DYN 1			DYN 2			DYN 3			DYN 4		
	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}
Orthosteric P3	200	-15.3	-5.57	180	-17.9	-8.25				350	-19.9	-7.54
Pocket A1	40	-13.7	-4.04				150	-12.6	-4.30	150	-19.2	-5.49
Pocket A4	190	-13.9	-4.52									
	60	-15.9	-4.15									
Pocket A6	50	-18.8	-5.57	110	-22.6	-7.56	170	-14.6	-4.07			

Averages				
Orthosteric P3	183	-13.3	-5.34	3 /4 dyn
Pocket A1	85	-11.4	-3.46	3 /4 dyn
Pocket A4	63	-7.5	-2.17	2 /4 dyn
Pocket A6	83	-14.0	-4.30	3 /4 dyn

Bcl_xL - Fragment p1

	DYN 1			DYN 2			DYN 3			DYN 4		
	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}
Orthosteric P3	30	-19.9	-6.05	50	-18.5	-5.39	70	-15.3	-6.37	270	-17.9	-5.18
Orthosteric P2	40	-14.6	-6.13	60	-15	-6.96	30	-13.6	-4.38	380	-14.9	-6.04
Pocket B1	80	-17.4	-5.51	360	-23.4	-6.13	60	-30.2	-6.41			
Pocket B2				90	-19.4	-4.46						
Pocket B3				20	-13.7	-7.24						
Pocket B4				35	-14	-5.3	50	-13.8	-5.53			
Pocket B5							25	-18.3	-6.67			
Pocket B6							60	-18.8	-5.83			
Pocket B7							20	-17.5	-5.14	20	-22.9	-3.77

Averages				
Orthosteric P3	105	-17.9	-5.75	4 /4 dyn
Orthosteric P2	128	-14.5	-5.88	4 /4 dyn
Pocket B1	125	-17.8	-4.51	3 /4 dyn
Pocket B2	23	-4.9	-1.12	1 /4 dyn
Pocket B3	5	-3.4	-1.81	1 /4 dyn
Pocket B4	21	-7.0	-2.71	2 /4 dyn
Pocket B5	6	-4.6	-1.67	1 /4 dyn
Pocket B6	15	-4.7	-1.46	1 /4 dyn
Pocket B7	10	-10.1	-2.23	2 /4 dyn

Bcl_xL - Fragment p2

	DYN 1			DYN 2			DYN 3			DYN 4		
	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}
Orthosteric P3	30	-18.3	-6.29				80	-19.7	-7.63	20	-19	-6.94
Pocket B1	30	-25.2	-7.11	100	-19.1	-5.58						
Pocket B5	20	-19.8	-6.12				240	-24.2	-6.24	60	-16.8	-5.01
Pocket B6							10	-18.6	-7.05	30	-17.4	-5.73
Pocket B8	160	-25.2	-5.62	20	-15.8	-5.13	75	-17.7	-5.93	80	-16.7	-6.05
Pocket B9	50	-24.1	-5.63	20	-16.7	-5.15	150	-23.6	-6.45			
Pocket B10							25	-13.8	-5.38			

Averages				
Orthosteric P3	33	-14.3	-5.22	3 /4 dyn
Pocket B1	33	-11.1	-3.17	2 /4 dyn
Pocket B5	80	-15.2	-4.34	3 /4 dyn
Pocket B6	10	-9.0	-3.20	2 /4 dyn
Pocket B8	84	-18.9	-5.68	4 /4 dyn
Pocket B9	55	-16.1	-4.31	3 /4 dyn
Pocket B10	6	-3.5	-1.35	1 /4 dyn

Bcl_xL - Fragment p3

	DYN 1			DYN 2			DYN 3			DYN 4		
	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}
Orthosteric P3	95	-21.9	-7.24	15	-15	-6.1	110	-16.6	-7.01	30	-23.4	-7.45
Pocket B4				10	-13.7	-6.43	65	-16.3	-7.81			
Pocket B5	30	-18.5	-6.25				30	-17.7	-5.44			
Pocket B6				40	-18.8	-7.08				40	-17.8	-6.98
Pocket B9				20	-10.9	-5.08						
Pocket B11	160	-17.8	-5.51				40	-19.6	-4.12			
Pocket B12										35	-26.4	-6.22

Averages				
Orthosteric P3	63	-19.2	-6.95	4 /4 dyn
Pocket B4	19	-7.5	-3.56	2 /4 dyn
Pocket B5	15	-9.1	-2.92	2 /4 dyn
Pocket B6	20	-9.2	-3.52	2 /4 dyn
Pocket B9	5	-2.7	-1.27	1 /4 dyn
Pocket B11	50	-9.4	-2.41	2 /4 dyn
Pocket B12	9	-6.6	-1.56	1 /4 dyn

Bcl_xL - Fragment p4

	DYN 1			DYN 2			DYN 3			DYN 4		
	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}
Orthosteric P2										40	-12.9	-4.03
Orthosteric P3										120	-15	-4.62
Pocket B1							10	-22.2	-6.12			
Pocket B6										130	-14.1	-4.29
Pocket B8				30	-20.4	-6.27						

Averages				
Orthosteric P2	10	-3.2	-1.01	1 /4 dyn
Orthosteric P3	30	-3.8	-1.16	1 /4 dyn
Pocket B1	3	-5.6	-1.53	1 /4 dyn
Pocket B6	33	-3.5	-1.07	1 /4 dyn
Pocket B8	8	-5.1	-1.57	1 /4 dyn

Bcl_xL - Fragment pd1

	DYN 1			DYN 2			DYN 3			DYN 4		
	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}
Pocket B2				20	-12.8	-3.98						
Pocket B4	90	-11.6	-6.99	110	-11.4	-6.37						
Pocket B6				30	-14.3	-5.97				30	-11.1	-6.74
Pocket B8							90	-19.3	-5.76			
Pocket B10							60	-11.9	-5.14			
Pocket B11							185	-13.7	-4.99			

Averages				
Pocket B2	5	-3.2	-1.00	1 /4 dyn
Pocket B4	50	-5.8	-3.34	2 /4 dyn
Pocket B6	15	-6.4	-3.18	2 /4 dyn
Pocket B8	23	-4.8	-1.44	1 /4 dyn
Pocket B10	15	-3.0	-1.29	1 /4 dyn
Pocket B11	46	-3.4	-1.25	1 /4 dyn

Bcl_xL - Fragment pd2

	DYN 1			DYN 2			DYN 3			DYN 4		
	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}	RT	MMGBSA	K_{DEEP}
Pocket B7							30	-15.3	-4.87	150	-24.8	-4.95
Pocket B8	100	-11.8	-4.58	20	-17.4	-5.54						
Pocket B10										90	-11.8	-6.05

Averages				
Pocket B7	45	-10.0	-2.46	2 /4 dyn
Pocket B8	30	-7.3	-2.53	2 /4 dyn
Pocket B10	23	-3.0	-1.51	1 /4 dyn